

# Genome-wide DNA methylation and expression patterns of microRNAs in relation to breast cancer subtypes among American women of African and European ancestry

Zhihong Gong (✉ [zhihong.gong@roswellpark.org](mailto:zhihong.gong@roswellpark.org))

Roswell Park Cancer Institute Department of Cancer Prevention and Population Sciences

<https://orcid.org/0000-0003-1129-1774>

Jianhong Chen

Roswell Park Comprehensive Cancer Center

Jie Wang

Roswell Park Comprehensive Cancer Center

Song Liu

Roswell Park Comprehensive Cancer Center

Christine B. Ambrosone

Roswell Park Comprehensive Cancer Center

Michael J. Higgins

Roswell Park Comprehensive Cancer Center

---

## Research article

**Keywords:** DNA methylation, microRNAs, breast cancer, African and European ancestry

**Posted Date:** September 1st, 2020

**DOI:** <https://doi.org/10.21203/rs.3.rs-66894/v1>

**License:**  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

# Abstract

**Background:** Aggressive high-grade, estrogen receptor negative (ER-) breast cancer is more common among American women of African ancestry (AA) than those of European ancestry (EA). The reasons remain largely unknown. Epigenetic mechanisms, particularly DNA methylation and altered microRNA (miRNA) expression, may contribute to racial differences in breast cancer. However, few studies have specifically characterized genome-wide DNA methylation-based modifications at the miRNA level in relation to ER+ and ER- breast cancer, and their functional role in the regulation of miRNA expression, especially among high risk AA women.

**Methods:** In this study, genome-wide DNA methylation and miRNA expression profiling was performed using the Illumina Infinium HumanMethylation450 Bead Chip platform and miRNA sequencing (miRNA-seq) in breast tumors from both AA and EA women.

**Results:** The genome-wide methylation screen identified a total of 7,191 unique CpGs mapped to 1,292 miRNA genes, which correspond to 2,035 unique mature miRNAs. Cluster analysis of these miRNA-associated methylation loci showed a clear pattern of ER-subtype differences. We identified differentially methylated loci (DMLs:  $(|\Delta \beta|) > 0.10$ ,  $FDR < 0.05$ ) between ER- and ER+ tumor subtypes, including 290 DMLs shared in both races, 317 and 136 were specific to AA and EA women, respectively. Integrated analysis of DNA methylation and corresponding miRNA expression identified certain DMLs whose methylation levels were significantly correlated with the expression of relevant miRNAs, such as multiple CpGs within *miR-190b* and *miR-135b* highly negatively correlated with their expression. Further target prediction and pathway analysis showed that these DNA methylation-dysregulated miRNAs are involved in multiple cancer-related pathways, including cell cycle G1-S growth factor, cytoskeleton remodeling, angiogenesis, EMT, and others such as signal transduction TGF- $\beta$ , Wnt, NOTCH, and ESR1-mediated signaling pathways.

**Conclusions:** Our results suggest that DNA methylation changes within miRNA genes are associated with altered miRNA expression, which may contribute to the network of subtype- and race-related tumor biological differences in breast cancer. These findings shed light on the epigenetic regulation of miRNA expression and provide insights into the relations of clinical-relevant miRNAs to their target genes and to serve as potential preventative and therapeutic targets.

## Introduction

Breast cancer is a heterogeneous disease, comprised of different clinical subtypes linked to disparate prognosis. However, as proposed by Anderson and colleagues (1), there are two primary etiologic subtypes, each associated with a distinct set of risk factors and genetic profiles that essentially correspond to estrogen receptor positive (ER+) and negative (ER-) disease, as an important marker for treatment options and prognosis. Compared to women diagnosed with ER+ breast cancer, those with ER- tumors in general have a poor prognosis, partly because of their aggressive phenotype and the lack of

targeted therapy. High-grade, ER- breast cancer is more common among American women of African ancestry (AA) than those of European ancestry (EA). However, much less is known about the underlying causes of increased risk of ER- breast cancer in AA women.

Changes in DNA methylation have been recognized as one of the most common molecular alterations in cancer (2). Aberrant methylation patterns have been frequently reported in breast cancer, with studies primarily focused on promoter regions of protein-coding genes (PCGs) (3–6). Specific methylation patterns have been associated with high tumor grade and ER negativity of breast cancer (4, 7–9). In our own studies and others, analyses restricted to CpG dinucleotides (CpGs) within and nearby PCGs, show differences in DNA methylation patterns in tumors by ER status and between AA and EA women (10–13). These results suggest that aggressive breast cancer subtypes may be related to altered gene methylation, and that methylation patterns may differ by ancestry.

Although it is well known that more than 75% of the genome is actively transcribed, yielding a large number of non-coding RNAs (ncRNAs), the genome-wide methylation patterns of ncRNAs in breast cancer remain largely unknown. MicroRNAs (miRNAs) are a well-characterized class of small ncRNAs that are key regulators of PCG expression, through induction of mRNA degradation or interference with mRNA translation (14). It is estimated that miRNAs regulate more than 50 percent of human genes (15) and are found to play an essential role in the regulation of many biologic processes in the development of cancer, including cell differentiation, apoptosis, immunity, and proliferation (16). Specifically in breast cancer, miRNA expression patterns in tumors have been linked to pathologic presentation and tumor subtypes (17–19), and specific miRNAs have been found to play essential roles in breast cancer invasion and metastasis (18, 20, 21). Moreover, in a recent study, we observed differential miRNA expression patterns by ER status and between races (22).

As current research has largely focused on the regulatory roles of miRNAs in cancer, their own regulation is not fully understood. Recent studies have shown that, similar to PCGs, miRNA encoding genes can be regulated by aberrant DNA methylation, and hence contribute to carcinogenesis (23, 24). Given that a single miRNA can regulate expression of multiple target genes, alteration of methylation patterns of miRNAs may have a much broader effect than at other loci. A number of studies have reported altered methylation of miRNA genes in breast cancer, however, most of these studies have been limited by one or more key factors, such as focusing on only a small number of candidate miRNA genes, few studies on human breast tumor tissues, and lack of data on direct correlation between methylation and expression levels (25). Furthermore, there are no studies that have specifically evaluated genome-wide methylation patterns of miRNAs in relation to differences in ER + and ER- breast cancer tissues from both AA and EA women. Taken together, identification and quantification of DNA methylation-based characteristics at the miRNA level in ER + and ER- tumors may improve our understanding of the molecular basis for the phenotypic differences between these two primary subtypes of breast cancer. In addition, examining DNA methylation changes in aggressive ER- tumor commonly observed in AA women may provide insights into underlying causes of breast cancer racial disparities.

In the current study, we characterized genome-wide DNA methylation patterns of CpGs within loci encoding miRNAs in breast tissues from both AA and EA women. We examined DNA methylation alteration patterns in ER+ and ER- tumors for both races and compared the variability within and across race. For a subgroup of tumors, we previously performed genome-wide miRNA expression profiling using next-generation sequencing, which allowed us to directly integrate DNA methylation and miRNA expression data to examine whether DNA methylation was associated with altered gene expression of corresponding miRNAs.

## **Materials And Methods**

### **Tissue samples and DNA methylation assay**

As described previously, we conducted genome-wide DNA methylation profiling of fresh-frozen breast tumor samples from 58 AA and 80 EA women using the Illumina Infinium HumanMethylation450 Bead Chip platform, with subsequent analyses restricted to CpG loci within or near PCGs (10). In the current analysis of these same samples, we focused on CpGs within loci encoding miRNAs. Briefly, breast tumor tissues were collected from patients who were treated at Roswell Park Comprehensive Cancer Center. Genomic DNA was isolated from banked specimens and stored at -80 °C until use. All samples were linked with detailed clinical information by the Roswell Park Biomedical Data Science Shared Resource, with tumor characteristics presented in Table 1.

Table 1  
 Characteristics of patients with breast cancer in current study

Factors	AA (n = 58, %)	EA (n = 80, %)
Age		
<50	18 (32)	22 (28)
50–68	20 (34)	29 (36)
>68	20 (34)	29 (36)
Estrogen Receptor Status		
Negative	26 (45)	25 (31)
Positive	32 (55)	55 (69)
Progesterone Receptor Status		
Negative	31 (53)	41 (51)
Positive	27 (47)	39 (49)
HER2 Status <sup>1</sup>		
Negative	47 (87)	44 (59)
Weak	1 (< 1)	4 (3)
Strong	7 (13)	26 (48)
Histological Grade <sup>2</sup>		
I (well differentiated)	0	2 (3)
II (moderately differentiated)	13 (23)	8 (10)
III (poorly differentiated)	44 (77)	68 (87)
<sup>1</sup> HER-2 status missing for 3 AA and 6 EA women;		
<sup>2</sup> Histologic grade data missing for 1 AA and 2 EA women		

Genome-wide DNA methylation analysis was carried out at Roswell Park Genetics Shared Resource using the Illumina Infinium HumanMethylation450 BeadChip platform, which interrogates > 485,000 CpG dinucleotides per sample at single-nucleotide resolution and covers 99% of RefSeq genes. In order to minimize the impact of batch effects, DNA samples from tumors were randomized on plates according to age, ancestry, and ER status. Following bisulfite treatment of DNA, subsequent steps for the methylation 450 BeadChip assay were carried out according to the manufacturer's instructions.

#### **DNA methylation data processing and analysis**

Hybridized and processed arrays are scanned, and the raw intensity is then extracted and processed by the minfi R package. We applied rigorous sample and locus specific quality control criteria, SWAN normalization, and correction for batch effects (26). We removed samples with poor detection p-values using the IMA package (27). Probes that were ambiguously mapped and those shown to contain SNPs were also excluded from the analysis (28, 29), leaving the final dataset contained 276,108 CpG loci.

A CpG locus to be associated with miRNA coding regions was defined as those located within 5Kb window of a miRNA gene. The unsupervised consensus clustering was performed on all unique miRNA-associated probes to examine overall methylation patterns in breast tumors, using the ConsensusClusterPlus R package (Bioconductor) with partitioning around medoids algorithm. The final consensus clustering was given by combining the ward.D linkage method and the Pearson correlation-based distance measurement. The Wilcoxon rank-sum test was used to evaluate the statistical significance for each probe in each of the comparisons. To adjust for multiple comparisons, the false discovery rate (FDR) was computed using the Benjamini and Hochberg approach. Differentially methylated loci (DML) were defined as CpGs with an absolute mean  $\beta$ -value difference ( $|\Delta \beta|$ ) at least 0.10 between subgroups and *FDR*-adjusted  $p < 0.05$ .

### **miRNA expression profiling and data processing**

Among breast tumors included in the DNA methylation assay, miRNA-seq was performed on a subset of tumors, including 58 tumors (29 AA and 29 EA women). As described previously (22), 1  $\mu$ g RNA was used for small RNA cDNA library preparation and then samples were sequenced on the Illumina HiSeq 2500 platform using high output 50-cycle single read sequencing. The miRNA quantification data were generated including expression levels for the known pre-miRNAs and mature miRNAs for all samples.

### **The Cancer Genome Atlas (TCGA) data processing and analysis**

To validate our findings, TCGA DNA methylation and miRNA expression data were downloaded from GDAC Broad Institute ([http://gdac.broadinstitute.org/runs/stddata\\_2016\\_01\\_28/data/BRCA/20160128/](http://gdac.broadinstitute.org/runs/stddata_2016_01_28/data/BRCA/20160128/)). Data generated using the same platform as the current study were included in this analysis. Patient and breast cancer clinical features were retrieved, such as age, race, and ER status. There were 537 TCGA samples with DNA methylation and miRNA expression data for the current analysis, with tumor characteristics described in Supplementary Table S1. All data were processed and analyzed following the same pipeline as used in the current study (above).

### **DNA methylation status and association with miRNA expression**

To examine whether DNA methylation status is associated with expression of corresponding miRNAs, we paired each CpG site with the corresponding miRNA, defined as probes located within 5Kb of a miRNA gene, as one CpG-miR Pair. Unique CpG-miR pairs were identified, in which one miRNA can be paired with multiple CpG sites, and similarly one CpG site can be paired with multiple miRNAs. Spearman's nonparametric correlation analysis was used to compute the correlation coefficient ( $\rho$ ) between methylation levels (beta values) and miRNA expression (log counts per million, logCPM) for each pair. Correlation analyses were carried out separately for tumors from AA and in EA women.

## miRNA target gene prediction and pathway analysis

For several top candidate CpG methylation-correlated-miRNAs, we further performed target prediction and pathway analysis using MetaCore (Clarivate Analytics), an integrated knowledge-based platform for comprehensive functional analysis of OMICs data and gene list. This method is based on a high-quality, manually curated database of molecular interactions, molecular pathways, and gene-disease associations. Specifically, in this analysis, we identified target genes of these miRNAs and their enriched biological processes and pathways within gene sets. MetaCore employs a hypergeometric model to determine significance of functional relationships. Functions with *FDR*-adjusted  $p < 0.05$  were considered significant.

## Results

### DNA methylation alterations at miRNA-associated loci in breast tumor tissue from AA and EA women

Based on the criterion that a CpG locus to be associated with a miRNA gene was defined as those located within 5Kb window of a given miRNA locus, a total of 7,191 unique CpGs were identified. These CpGs were mapped to 1,292 miRNA genes, which correspond to 2,035 unique mature miRNAs based in the latest release of miRBase version 22 (<http://www.mirbase.org/>). As shown in Fig. 1, unsupervised consensus clustering analysis of the methylation levels at these miRNA-associated CpG loci identified two major clusters. Cluster 1 contains primarily ER + breast tumors; cluster 2 is enriched for ER- tumors.

As no known studies have specifically examined genome-wide DNA methylation patterns of miRNA encoding genes in breast tumors from AA populations and there is a clear pattern of ER-subtype differences (Fig. 1), in the current study, we focused on analyses to identify differentially methylated loci (DML) between ER- and ER + tumors in AA and EA women, separately. As shown in Fig. 2A-B, we identified 607 DMLs in tumors from AA women, while there were 426 DMLs in tumors from EA women. Among DMLs identified in ER- versus ER + tumors from either AA or EA women, 290 DMLs (146 hyper- and 144 hypo- methylated) were identified in both races, 317 were specific to AAs (198 hyper- and 119 hypo- methylated), and 136 were specific to EAs (34 hyper- and 102 hypo- methylated). We further presented these DMLs in the volcano plot for AA women (Fig. 2C) and for EA women (Fig. 2D), which clearly demonstrated a significantly higher frequency of hypermethylated loci in tumors from AA women compared to EA women. The full list of these significant DMLs by ER subtype for each group is shown in Supplemental Table S2.

### Validation of DMLs in the TCGA dataset

We validated DNA methylation for DMLs identified above using data from TCGA, which included 141 AA and 396 EA cases in this analysis. Data on 537 (out of 607 DMLs in AAs) and 385 (out of 426 DMLs in EAs) CpGs were available in the TCGA dataset. As shown in Supplemental Figure S1, we observed a high concordance between DMLs detected in our data and those identified using the TCGA dataset, i.e., all probes showed the same direction of methylation change by ER status, although in TCGA dataset,

methylation changes of 16 DMLs in AAs and the 3 DMLs in EAs did not reach the statistical significance (dots in red).

### **Impact of aberrant DNA methylation on miRNA expression**

To examine whether these DMLs are associated with expression of their corresponding miRNAs, we paired each CpG with the corresponding miRNA defined as within 5Kb window, with a total of 1,408 unique CpG-miR pairs identified in our data. In cancers from AA women, we identified 224 (94 negative and 130 positive) significant correlations between methylation and expression levels, involving 136 unique CpGs at which differential methylation was associated with expression of 124 mature miRNAs ( $p < 0.05$ ) from 97 unique miRNA genes. In tumors from EA women, we found 97 (45 negative and 52 positive) correlations, corresponding to 70 CpGs and 64 mature miRNAs ( $p < 0.05$ ) from 51 unique miRNA genes. The correlation coefficients for each of these significant CpG-miR pairs are shown in Fig. 3A, with detailed information on the CpG loci, mature miRNAs, and their correlation coefficients listed in Table S3. Notably, significant negative correlations for multiple CpG loci were observed for *miR-190b* and *miR-135b* in both race groups, whereas positive correlations were observed for members of the *miR-224/miR-452* cluster. As shown in Fig. 3B, the scatter plots showed the correlation patterns of several methylation loci with the corresponding miRNA expression on the *miR-190b*, *miR-135b*, and *miR-452/-224*. Results suggest that DNA methylation changes appeared to drive miRNA expression alterations and that these correlations were somewhat stronger in AAs than in EAs.

We further validated these findings in TCGA. The majority of correlations identified in our study showed consistent directions in TCGA, with detailed information listed in supplemental Table S4. In Figure S2, the scatter plots showed significant correlations for multiple CpGs with their corresponding miRNAs, i.e., the *miR-190b*, *miR-135b*, *miR-224/-452*. As shown, the methylation level of a CpG mapped within the *miR-190b* locus was negatively correlated with expression levels of the mature *miR-190b-5p*, in which high methylation levels were associated with low expression levels in both our and TCGA data. Similar findings were observed for multiple probes of *miR-135b*. These results further confirmed that DNA methylation modifications in miRNA encoding genes play regulatory role of expression of these miRNAs.

### **miRNA target genes and pathway analysis**

To investigate the biological relevance of these findings, we performed target prediction and pathway enrichment analysis using MetaCore, focusing on several top candidate miRNAs at which DNA methylation was highly correlated with miRNA expression, including *miR-135b-5p*, *miR-190b-5p*, and *miR-224-5p/miR-452-3p*. Details on canonical pathways, molecular process networks, and targets for each miRNA are presented in supplemental Table S5. These miRNAs deregulated by DNA methylation in this setting are involved in multiple cancer-related pathways, indicating their role in various key molecular processes related to and beyond ER biology. As shown in Fig. 4, top enriched pathway and process networks related to *miR-190b-5p* and *miR-135b-5p* include cell cycle G1-S growth factor or interleukin regulations, cytoskeleton remodeling, angiogenesis, EMT, and others such as signal transduction TGF- $\beta$ , Wnt, NOTCH, and ESR1-mediated signaling pathways. Targets from these process networks were

identified, including various transcription factors and protein kinase, such as *ESR1*, *GSK3 $\beta$* , *MMP-2*, *MMP-9*, *TGF- $\beta$* , *mTOR*, and *PI3KCA*.

## Discussion

DNA methylation has been reported as a mechanism that may cause dysregulation of mature miRNAs and consequently associated with breast cancer risk and progression. However, few studies have specifically examined this epigenetic modification at miRNA level in relation to ER+ and ER- breast cancer, and its functional role in the regulation of miRNA expression, especially in AA women who are at high risk of aggressive, ER- disease. In this study, we evaluated methylation patterns of miRNA genes and their effect on miRNA expression in breast tumors from both AA and EA women. The genome-wide methylation screen identified a set of DNA methylation loci within 5Kb of miRNA genes that were differentially methylated between ER- and ER+ tumor subtypes in tumors from both races, or specific to AA or EA women. Integrated analysis of DNA methylation and miRNA expression further identified certain DMLs whose methylation levels were significantly correlated with the expression of relevant miRNAs. Thus, DNA methylation changes associated with altered miRNA expression may contribute to the network of molecular differences of ER subtypes of breast cancer and may differ by race.

We identified many DMLs that can significantly differentiate ER- and ER+ tumor subtypes. Some of the top DMLs mapped to, or were in the vicinity of, miRNAs that are differentially expressed in ER- versus ER+ tumors, such as multiple CpGs in *miR-135b*, *-190b*, and *-224*, as shown in our previous study and in other studies (22, 30). Many other DMLs were novel relative to ER subtypes, such as CpGs associated with *miR-125b1*, *-2053*, *-3132*, and *-4736*. Some of these identified DMLs were common in both races and others were specific to AA or EA women. We also noted that there were more DMLs by ER subtypes in AA tumors compared to EA tumors. Moreover, while ER-subtype DMLs common to both races showed similar proportions of hyper- and hypo-methylated loci, interestingly, we noted significant differences in proportions for race-specific DMLs. We observed a higher frequency of hypermethylated miRNA genes in ER- compared to ER+ tumors from AA women, but a lower frequency of hypermethylation in ER- versus ER+ tumors from EA women. Consistent with our findings for tumors from EA women, several studies focused on DNA methylation of coding genes observed a higher frequency of hypermethylated genes in ER+ tumors (8, 31, 32). One explanation for this may be that ER-related signaling molecules may enhance the activity of DNA methylation transferases to catalyze the transfer of a methyl group to a DNA segment, thus lead to the specific gain of DNA methylation in ER+ tumors. By contrast, we observed in AA women that there were more DMLs associated with ER tumor subtypes and that there was a higher frequency of hypermethylated miRNA loci in ER- tumors, which warrant further investigations. Nevertheless, the majority of these DMLs were confirmed in TCGA data, suggesting a differential DNA methylation pattern of miRNA encoding genes relevant to breast cancer subtypes and race.

Another goal of our work was to identify DNA methylation-associated changes in miRNA expression that may be involved in phenotypic differences in breast cancer subtypes in both AA and EA women. Leveraging the availability of both genome-wide DNA methylation and miRNA expression data from the

same patients, we were able to directly link both types of data and examined whether altered DNA methylation at specific CpG loci was associated with the expression of neighboring miRNAs. As expected, we observed that altered methylation at certain DMLs was negatively correlated with expression of corresponding miRNAs in breast tumors of both AA and EA women. Several top correlated CpG-miR pairs include multiple loci within the *miR-190b* and *miR-135b*. *MiR-190b* has been reported to be highly upregulated in ER+ compared to ER- breast cancers from several other studies including our own (22, 33–35). In the present study, our results provide evidence that differences in *miR-190b* expression in breast cancer subtypes is associated with methylation alterations. This observation was consistent with results from our analysis of TCGA dataset and from a recently published study (35). Results from studies also suggest that expression of *miR-190b* in primary tumors may play a role with breast cancer recurrence, tamoxifen resistance, breast cancer and overall survival, but its exact role in carcinogenesis remains to elucidate (33, 34, 36). Interestingly, as low expression is associated with poor survival, we observed the lowest *miR-190b* expression in ER- tumors from AA women among all tumor-race subgroups, which may contribute to the poor survival observed in this population. By contrast, we observed CpG promoter hypomethylation and upregulation of *miR-135b* in ER- versus ER+ tumors. Consistent with these results, several studies have reported *miR-135b* was upregulated in ER- or triple negative breast cancers and its overexpression promotes breast cancer progression and metastasis through regulation of *ERα*, *AR*, and other genes such as *LATS2*, *TGFβ*, *WNT*, and *HIF1AN* (37–41). Taken together, confirmed in TCGA data, these consistent negative correlations between CpG methylation and miRNA expression support the notion that, at least in some cases, altered miRNA expression may occur through mechanisms involving loss or gain of promoter DNA methylation during early stages of carcinogenesis and contribute to the development of specific subtypes of the disease. These DNA methylation-dysregulated network of miRNA genes could be clinically relevant predictive markers and served as potential new therapeutic targets of anticancer therapy. It may be possible to restore the functionality of the dysregulated miRNA network through epigenetic drugs. Future study is necessary to fully understand the mechanisms of underlying biological functions and their clinical implications.

Although most studies have primarily focused on the relationship of DNA methylation and their negative control of gene expression, we revealed positive correlations between methylation and expression for many CpG-miRNA pairs, similar to what has been consistently reported in cancer studies that investigate correlation patterns between CpG methylation and gene expression at the genome-wide scale (42). Notably, we found that DNA methylation of multiple loci within the *miR-224/miR-452* cluster were positively associated with miRNA expression in both races. These two intronic miRNAs are located within a locus on chromosome Xq28 encompassing *GABRE* gene. We observed hypermethylation at these CpG loci and increased *miR-224/miR-452* expression in ER- compared to ER+ tumors. Studies have reported both up- and down-regulation of *miR-224* in various cancers, suggesting its cell type-specific expression and function (43–45). Specifically, in breast cancer, and consistent with our study, one study reported increased *miR-224* expression in both triple negative breast cancer cell lines and tumor tissues, and that its knockdown in aggressive triple negative breast cancer cells reduced proliferation, migration and invasion (46). These results may suggest more diverse mechanisms of epigenetic regulation, which may

involve DNA methylation of insulator or repressor elements, transcription factor binding, chromatin and histone modifications during carcinogenesis. Such complexity warrants further investigation of these miRNAs and findings could provide implications for understanding the molecular basis for the phenotypic differences by breast cancer subtype and race.

We further performed target predictions and examined the major biological functions and signaling pathways of several significant methylation-dysregulated miRNAs. Various important metastasis- and cancer-related pathways were enriched, including cell cycle regulation, angiogenesis, EMT, and Wnt, TGF- $\beta$  and ESR1-mediated signaling. Specifically, we found that these miRNAs participate in these pathways through interactions with targets such as *ESR1*, *GSK3 $\beta$* , *MMP-2/-9*, *TGF $\beta$* , *mTOR*, *SMAD4*, *c-Myc*, *HIF1 $\alpha$* , and *PI3KCA*, some of which have been shown to regulate various signaling pathways and cellular functions in breast carcinogenesis (47–50). These results suggest that these miRNAs target many key genes and work as a module, contributing to breast cancer phenotypic differences and aggressiveness. Further experimental validation will be needed to confirm these findings.

## Conclusions

In summary, based on genome-wide profiling of tumor DNA methylation and microRNA expression, our results suggest that DNA methylation patterns in miRNA encoding genes differ between breast cancers according to cancer subtype and race, and that this altered methylation may affect miRNA expression. Further pathway analysis identified their potential role in modulating cancer-related key biological processes. These findings provide the basis for further functional analyses in experimental studies, which are likely to further inform the underpinning of subtype- and race-related tumor biological differences in breast cancer. Our study sheds light on the epigenetic regulation of miRNA expression and provide insights into the relations of clinical-relevant miRNAs to their target genes and to serve as potential preventative and therapeutic targets.

## Abbreviations

AA: African ancestry

CpGs: CpG dinucleotides

DML: Differentially methylated loci

ER: Estrogen receptor

EA: European ancestry

FDR: False discovery rate

miRNA: microRNA

ncRNAs: non-coding RNAs

PCGs: protein-coding genes

## Declarations

### *Ethics approval and consent to participate:*

The protocol was approved by the Roswell Park Institutional Review Board, with signed patient consent for use of remnant tissue for research.

### *Consent for publication:*

Not applicable

### *Availability of data and materials:*

The DNA methylation and miRNA expression data have been deposited in GEO.

All other data analyzed and presented in this study are in the Supplementary tables and figures.

### *Competing interest:*

Not applicable

### *Funding:*

This work was supported by the grants from the National Institutes of Health/National Cancer Institute (K07 CA178293, R01 CA1332641). The Roswell Park Pathology Network Shared Resource, Biomedical Data Science Shared Resource, Genetics Shared Resource, and Bioinformatics are Roswell Park CCSG Shared Resources (NIH P30 CA016056).

### *Authors' contributions:*

ZG, CBA, and MJH conceived of and designed the study. JC, JW, SL acquired and analyzed data. ZG drafted the manuscript. All authors revised the manuscript and read and approved the final manuscript.

## References

1. Anderson WF, Rosenberg PS, Prat A, Perou CM, Sherman ME. How many etiological subtypes of breast cancer: two, three, four, or more? *J Natl Cancer Inst.* 2014;106.
2. Wajed SA, Laird PW, DeMeester TR. DNA methylation: an alternative pathway to cancer. *Ann Surg.* 2001;234:10-20.

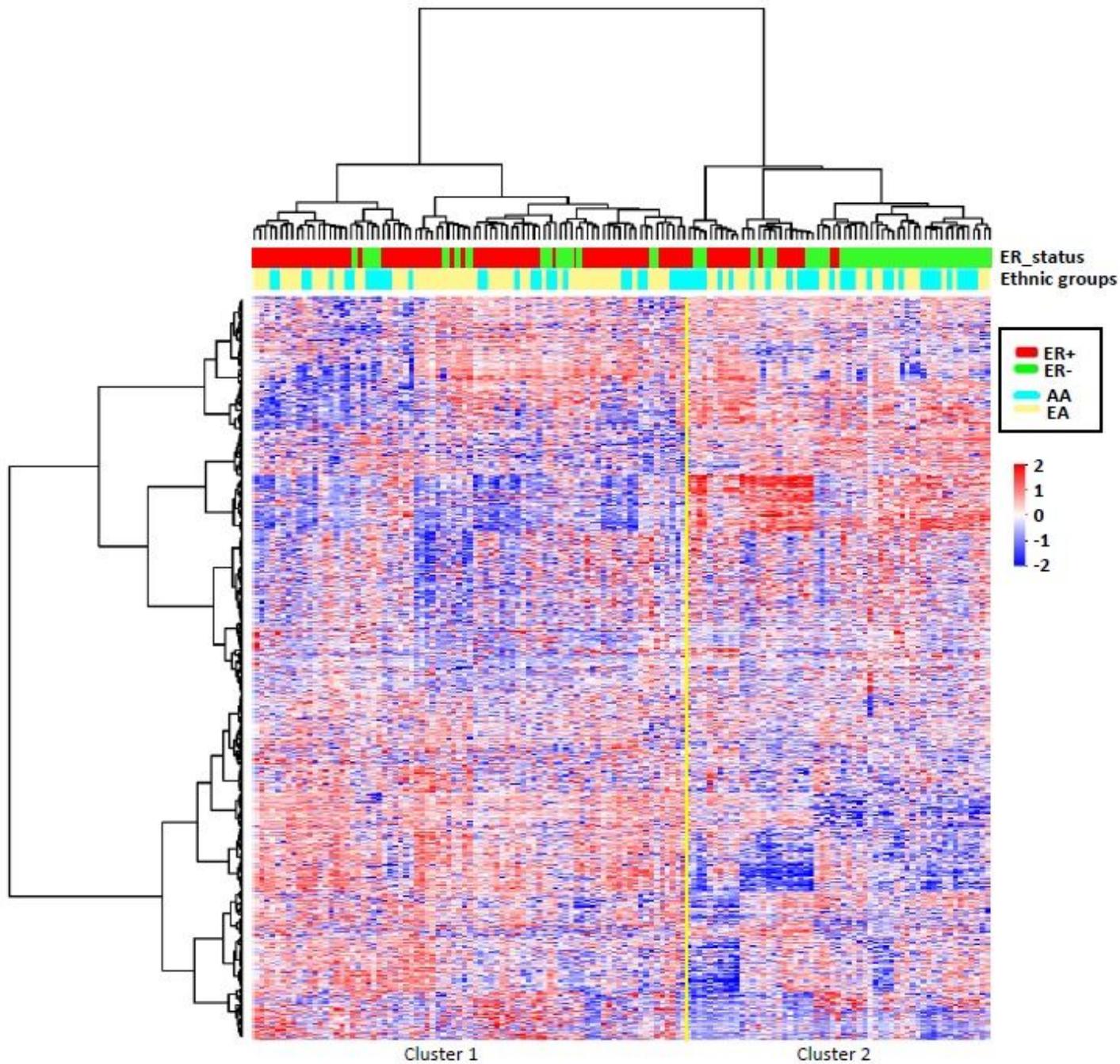
3. Paredes J, Albergaria A, Oliveira JT, Jeronimo C, Milanezi F, Schmitt FC. P-cadherin overexpression is an indicator of clinical outcome in invasive breast carcinomas and is associated with CDH3 promoter hypomethylation. *Clin Cancer Res.* 2005;11:5869-77.
4. Takahashi Y, Miyoshi Y, Takahata C, Irahara N, Taguchi T, Tamaki Y, Noguchi S. Down-regulation of LATS1 and LATS2 mRNA expression by promoter hypermethylation and its association with biologically aggressive phenotype in human breast cancers. *Clin Cancer Res.* 2005;11:1380-5.
5. Luo S, Chen J, Mo X. The association of PTEN hypermethylation and breast cancer: a meta-analysis. *Onco Targets Ther.* 2016;9:5643-50.
6. Yu P, Guo Y, Yusufu M, Liu Z, Wang S, Yin X, Peng G, Wang L, Zhao X, Guo H, Huang T, Liu C. Decreased expression of EZH2 reactivates RASSF2A by reversal of promoter methylation in breast cancer cells. *Cell Biol Int.* 2016;40:1062-70.
7. Bardowell SA, Parker J, Fan C, Crandell J, Perou CM, Swift-Scanlan T. Differential methylation relative to breast cancer subtype and matched normal tissue reveals distinct patterns. *Breast Cancer Res Treat.* 2013;142:365-80.
8. Fackler MJ, Umbricht CB, Williams D, Argani P, Cruz LA, Merino VF, Teo WW, Zhang Z, Huang P, Visvanathan K, Marks J, Ethier S, Gray JW, Wolff AC, Cope LM, Sukumar S. Genome-wide methylation analysis identifies genes specific to breast cancer hormone receptor status and risk of recurrence. *Cancer Res.* 2011;71:6195-207.
9. Li L, Lee KM, Han W, Choi JY, Lee JY, Kang GH, Park SK, Noh DY, Yoo KY, Kang D. Estrogen and progesterone receptor status affect genome-wide DNA methylation profile in breast cancer. *Hum Mol Genet.* 2010;19:4273-7.
10. Ambrosone CB, Young AC, Sucheston LE, Wang D, Yan L, Liu S, Tang L, Hu Q, Freudenheim JL, Shields PG, Morrison CD, Demissie K, Higgins MJ. Genome-wide methylation patterns provide insight into differences in breast tumor biology between American women of African and European ancestry. *Oncotarget.* 2014;5:237-48.
11. Espinal AC, Buas MF, Wang D, Cheng DT, Sucheston-Campbell L, Hu Q, Yan L, Payne-Ondracek R, Cortes E, Tang L, Gong Z, Zirpoli G, Khoury T, Yao S, Omilian A, Demissie K, Bandera EV, Liu S, Ambrosone CB, Higgins MJ. FOXA1 hypermethylation: link between parity and ER-negative breast cancer in African American women? *Breast Cancer Res Treat.* 2017;166:559-68.
12. Mehrotra J, Ganpat MM, Kanaan Y, Fackler MJ, McVeigh M, Lahti-Domenici J, Polyak K, Argani P, Naab T, Garrett E, Parmigiani G, Broome C, Sukumar S. Estrogen receptor/progesterone receptor-negative breast cancers of young African-American women have a higher frequency of methylation of multiple genes than those of Caucasian women. *Clin Cancer Res.* 2004;10:2052-7.
13. Wang S, Dorsey TH, Terunuma A, Kittles RA, Ambs S, Kwabi-Addo B. Relationship between tumor DNA methylation status and patient characteristics in African-American and European-American women with breast cancer. *PLoS One.* 2012;7:e37928.
14. He L, Hannon GJ. MicroRNAs: small RNAs with a big role in gene regulation. *Nat Rev Genet.* 2004;5(7):522-31. Epub 2004/06/24. doi: 10.1038/nrg1379. PubMed PMID: 15211354.

15. Bartel DP. MicroRNAs: target recognition and regulatory functions. *Cell*. 2009;136:215-33.
16. Bartel DP. MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell*. 2004;116:281-97.
17. Blenkiron C, Goldstein LD, Thorne NP, Spiteri I, Chin SF, Dunning MJ, Barbosa-Morais NL, Teschendorff AE, Green AR, Ellis IO, Tavare S, Caldas C, Miska EA. MicroRNA expression profiling of human breast cancer identifies new markers of tumor subtype. *Genome Biol*. 2007;8:R214.
18. Iorio MV, Ferracin M, Liu CG, Veronese A, Spizzo R, Sabbioni S, Magri E, Pedriali M, Fabbri M, Campiglio M, Menard S, Palazzo JP, Rosenberg A, Musiani P, Volinia S, Nenci I, Calin GA, Querzoli P, Negrini M, Croce CM. MicroRNA gene expression deregulation in human breast cancer. *Cancer Res*. 2005;65:7065-70.
19. Mattie MD, Benz CC, Bowers J, Sensinger K, Wong L, Scott GK, Fedele V, Ginzinger D, Getts R, Haqq C. Optimized high-throughput microRNA expression profiling provides novel biomarker assessment of clinical prostate and breast cancer biopsies. *Mol Cancer*. 2006;5:24.
20. Tavazoie SF, Alarcon C, Oskarsson T, Padua D, Wang Q, Bos PD, Gerald WL, Massague J. Endogenous human microRNAs that suppress breast cancer metastasis. *Nature*. 2008;451:147-52.
21. Valastyan S, Reinhardt F, Benaich N, Calogrias D, Szasz AM, Wang ZC, Brock JE, Richardson AL, Weinberg RA. A pleiotropically acting microRNA, miR-31, inhibits breast cancer metastasis. *Cell*. 2009;137:1032-46.
22. Gong Z, Wang J, Wang D, Buas MF, Ren X, Freudenheim JL, Belinsky SA, Liu S, Ambrosone CB, Higgins MJ. Differences in microRNA expression in breast cancer between women of African and European ancestry. *Carcinogenesis*. 2019;40:61-9.
23. Strmsek Z, Kunej T. MicroRNA Silencing by DNA Methylation in Human Cancer: a Literature Analysis. *Noncoding RNA*. 2015;1:44-52.
24. Suzuki H, Maruyama R, Yamamoto E, Kai M. DNA methylation and microRNA dysregulation in cancer. *Mol Oncol*. 2012;6:567-78.
25. Lehmann U. Aberrant DNA methylation of microRNA genes in human breast cancer - a critical appraisal. *Cell Tissue Res*. 2014;356:657-64.
26. Leek JT, Johnson WE, Parker HS, Jaffe AE, Storey JD. The sva package for removing batch effects and other unwanted variation in high-throughput experiments. *Bioinformatics*. 2012;28:882-3.
27. Wang D, Yan L, Hu Q, Sucheston LE, Higgins MJ, Ambrosone CB, Johnson CS, Smiraglia DJ, Liu S. IMA: an R package for high-throughput analysis of Illumina's 450K Infinium methylation data. *Bioinformatics*. 2012;28:729-30.
28. Chen YA, Lemire M, Choufani S, Butcher DT, Grafodatskaya D, Zanke BW, Gallinger S, Hudson TJ, Weksberg R. Discovery of cross-reactive probes and polymorphic CpGs in the Illumina Infinium HumanMethylation450 microarray. *Epigenetics*. 2013;8:203-9.
29. Zhang X, Mu W, Zhang W. On the analysis of the illumina 450k array data: probes ambiguously mapped to the human genome. *Front Genet*. 2012;3:73.

30. Howard EW, Yang X. microRNA Regulation in Estrogen Receptor-Positive Breast Cancer and Endocrine Therapy. *Biol Proced Online*. 2018;20:17.
31. Bediaga NG, Acha-Sagredo A, Guerra I, Viguri A, Albaina C, Ruiz Diaz I, Rezola R, Alberdi MJ, Dopazo J, Montaner D, Renobales M, Fernandez AF, Field JK, Fraga MF, Liloglou T, de Pancorbo MM. DNA methylation epigenotypes in breast cancer molecular subtypes. *Breast Cancer Res*. 2010;12:R77.
32. Williams KE, Jawale RM, Schneider SS, Otis CN, Pentecost BT, Arcaro KF. DNA methylation in breast cancers: Differences based on estrogen receptor status and recurrence. *J Cell Biochem*. 2019;120:738-55.
33. Cizeron-Clairac G, Lallemand F, Vacher S, Lidereau R, Bieche I, Callens C. MiR-190b, the highest up-regulated miRNA in ERalpha-positive compared to ERalpha-negative breast tumors, a new biomarker in breast cancers? *BMC Cancer*. 2015;15:499.
34. Dai W, He J, Zheng L, Bi M, Hu F, Chen M, Niu H, Yang J, Luo Y, Tang W, Sheng M. miR-148b-3p, miR-190b, and miR-429 Regulate Cell Progression and Act as Potential Biomarkers for Breast Cancer. *J Breast Cancer*. 2019;22:219-36.
35. Frick E, Gudjonsson T, Eyfjord J, Jonasson J, Tryggvadottir L, Stefansson O, Sigurdsson S. CpG promoter hypo-methylation and up-regulation of microRNA-190b in hormone receptor-positive breast cancer. *Oncotarget*. 2019;10:4664-78.
36. Lyng MB, Laenkholm AV, Sokilde R, Gravgaard KH, Litman T, Ditzel HJ. Global microRNA expression profiling of high-risk ER+ breast cancers from patients receiving adjuvant tamoxifen mono-therapy: a DBCG study. *PLoS One*. 2012;7:e36170.
37. Aakula A, Leivonen SK, Hintsanen P, Aittokallio T, Ceder Y, Borresen-Dale AL, Perala M, Ostling P, Kallioniemi O. MicroRNA-135b regulates ERalpha, AR and HIF1AN and affects breast and prostate cancer cell growth. *Mol Oncol*. 2015;9:1287-300.
38. Hua K, Jin J, Zhao J, Song J, Song H, Li D, Maskey N, Zhao B, Wu C, Xu H, Fang L. miR-135b, upregulated in breast cancer, promotes cell growth and disrupts the cell cycle by regulating LATS2. *Int J Oncol*. 2016;48:1997-2006.
39. Paszek S, Gablo N, Barnas E, Szybka M, Morawiec J, Kolacinska A, Zawlik I. Dysregulation of microRNAs in triple-negative breast cancer. *Ginekol Pol*. 2017;88:530-6.
40. Pu T, Shen M, Li S, Yang L, Gao H, Xiao L, Zhong X, Zheng H, Liu Y, Ye F, Bu H. Repression of miR-135b-5p promotes metastasis of early-stage breast cancer by regulating downstream target SDCBP. *Lab Invest*. 2019;99:1296-308.
41. Uva P, Cossu-Rocca P, Loi F, Pira G, Murgia L, Orru S, Floris M, Mureni MR, Sanges F, Carru C, Angius A, De Miglio MR. miRNA-135b Contributes to Triple Negative Breast Cancer Molecular Heterogeneity: Different Expression Profile in Basal-like Versus non-Basal-like Phenotypes. *Int J Med Sci*. 2018;15:536-48.
42. Long MD, Smiraglia DJ, Campbell MJ. The Genomic Impact of DNA CpG Methylation on Gene Expression; Relationships in Prostate Cancer. *Biomolecules*. 2017;7.

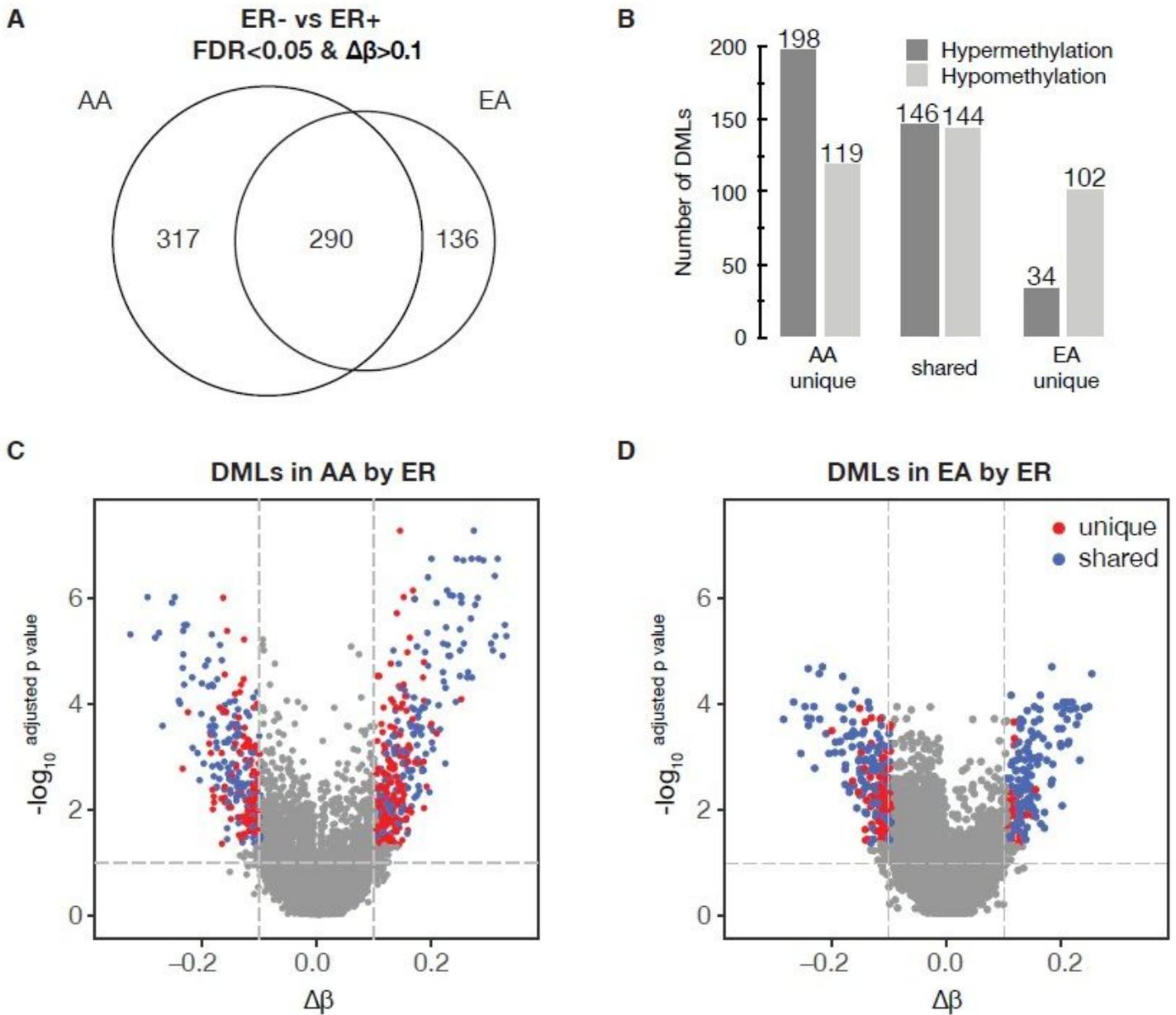
43. Fassan M, Cui R, Gasparini P, Mescoli C, Guzzardo V, Vicentini C, Munari G, Loupakis F, Lonardi S, Braconi C, Scarpa M, D'Angelo E, Pucciarelli S, Angriman I, Agostini M, D'Inca R, Farinati F, Gafa R, Lanza G, Frankel WL, Croce CM, Valeri N, Rugge M. miR-224 Is Significantly Upregulated and Targets Caspase-3 and Caspase-7 During Colorectal Carcinogenesis. *Transl Oncol.* 2019;12:282-91.
44. Gan BL, Zhang LJ, Gao L, Ma FC, He RQ, Chen G, Ma J, Zhong JC, Hu XH. Downregulation of miR2245p in prostate cancer and its relevant molecular mechanism via TCGA, GEO database and in silico analyses. *Oncol Rep.* 2018;40:3171-88.
45. Li S, Zhang J, Zhao Y, Wang F, Chen Y, Fei X. miR-224 enhances invasion and metastasis by targeting HOXD10 in non-small cell lung cancer cells. *Oncol Lett.* 2018;15:7069-75.
46. Zhang L, Zhang X, Wang X, He M, Qiao S. MicroRNA-224 Promotes Tumorigenesis through Downregulation of Caspase-9 in Triple-Negative Breast Cancer. *Dis Markers.* 2019;2019:7378967.
47. Paplomata E, O'Regan R. The PI3K/AKT/mTOR pathway in breast cancer: targets, trials and biomarkers. *Ther Adv Med Oncol.* 2014;6:154-66.
48. Gomes LR, Terra LF, Wailemann RA, Labriola L, Sogayar MC. TGF-beta1 modulates the homeostasis between MMPs and MMP inhibitors through p38 MAPK and ERK1/2 in highly invasive breast cancer cells. *BMC Cancer.* 2012;12:26.
49. Lee KM, Giltane JM, Balko JM, Schwarz LJ, Guerrero-Zotano AL, Hutchinson KE, Nixon MJ, Estrada MV, Sanchez V, Sanders ME, Lee T, Gomez H, Lluch A, Perez-Fidalgo JA, Wolf MM, Andrejeva G, Rathmell JC, Fesik SW, Arteaga CL. MYC and MCL1 Cooperatively Promote Chemotherapy-Resistant Breast Cancer Stem Cells via Regulation of Mitochondrial Oxidative Phosphorylation. *Cell Metab.* 2017;26:633-47 e7.
50. Vijay GV, Zhao N, Den Hollander P, Toneff MJ, Joseph R, Pietila M, Taube JH, Sarkar TR, Ramirez-Pena E, Werden SJ, Shariati M, Gao R, Sobieski M, Stephan CC, Sphyris N, Miura N, Davies P, Chang JT, Soundararajan R, Rosen JM, Mani SA. GSK3beta regulates epithelial-mesenchymal transition and cancer stem cell properties in triple-negative breast cancer. *Breast Cancer Res.* 2019;21:37.

## Figures



**Figure 1**

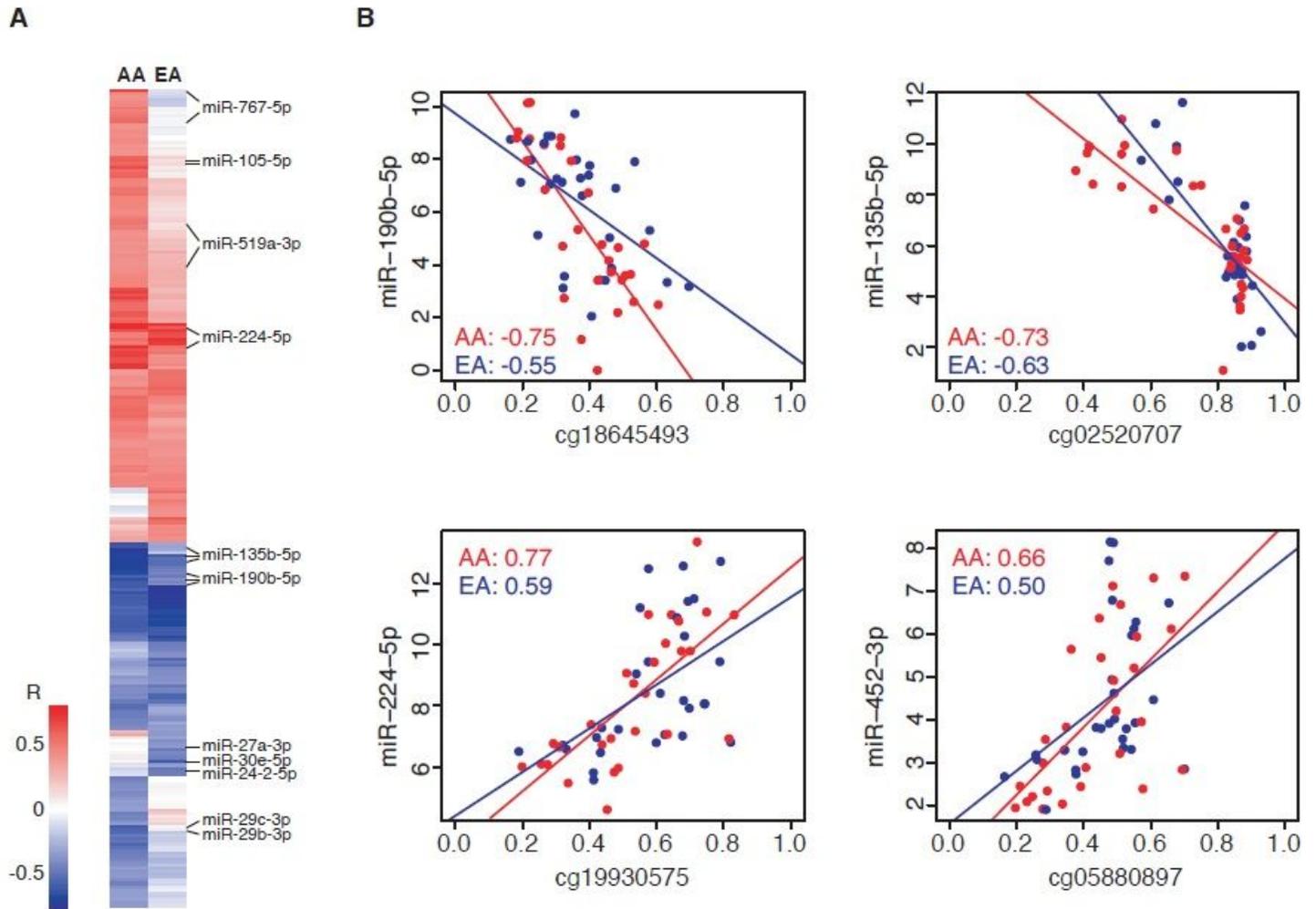
Unsupervised consensus cluster analysis of miRNA methylation patterns of breast tissues. Two major clusters were identified. Cluster 1 contains primarily ER+ breast tumors; cluster 2 is enriched for ER- tumors.



**Figure 2**

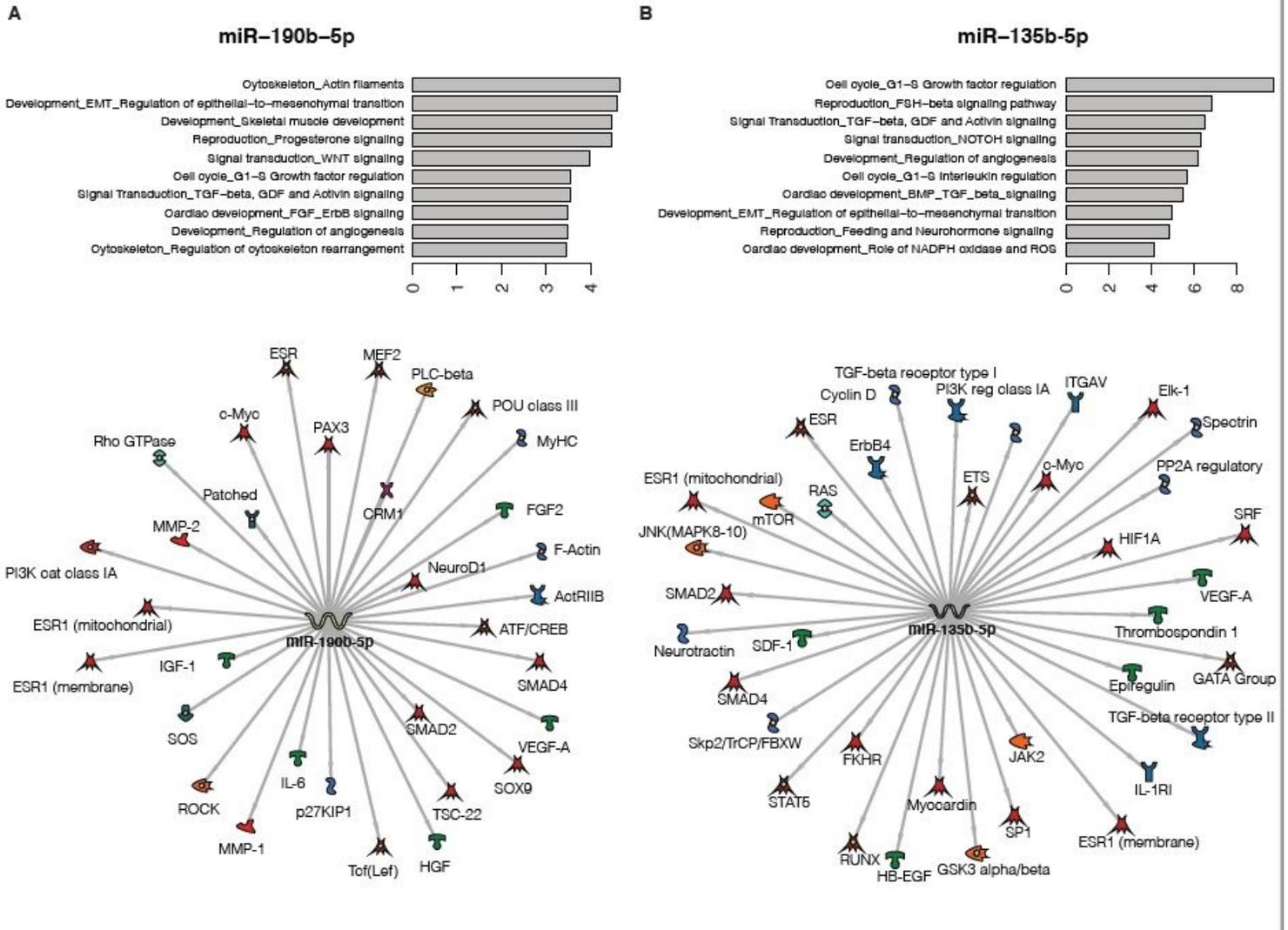
The differentially methylated loci (DML:  $|\Delta\beta| \geq 0.10$ ; FDR<0.05) between ER- and ER+ tumors in AA and EA women. (A) Venn diagram showing the number of DMLs by ER status specific to EA women (n=136), common in both AA and EA women (n=290), and specific to AA women (n=317). (B) The histogram shows the proportions of hyper- and hypo- methylated by ER-DMLs in AA-specific, AA-EA Shared, and EA-specific groups. (C and D) Volcano plots showing distributions of hyper- and hypo-methylated DMLs by ER status in AA (C) and in EA women (D). Dots plotted in red represent DMLs specific to either AA or EA women; Dots plotted in blue represent DMLs shared in both AA and EA women; red or blue dots in the right panel represent for hypermethylated and those in the left panel for hypomethylated DMLs. Dots plotted in grey are not DMLs in either AA or EA women. The mean

methylation difference ( $|\Delta \beta|$ ) in ER- versus ER+ tumors is on the X axis, and the Y axis shows the log 10 of FDR (adjusted P-value).



**Figure 3**

Correlation coefficients between methylation and miRNA expression levels on by ER-DMLs and their corresponding miRNAs. (A) The heatmap shows the correlation coefficients (R) between each CpG-miR pair in AA and EA tumors. Red bars represent positive correlation and blue bars represent negative correlation. (B) The Scatter plots relate the methylation and expression levels on selected top correlated CpG-miR pairs including specific probes with miR-190b-5p, miR-135b-5p, miR-224-5p, and miR-452-3p. For each CpG-miR pair, the methylation level (beta value) is on the X axis, and the expression level (log counts per million, logCPM) of corresponding miRNA is on Y axis.



**Figure 4**

Functional enrichment analysis on targets from miR-135b-5p and miR-190b-5p using MetaCore. (A) Target gene network and top ten pathway networks for miR-190b-5p. (B) Target gene network and top ten pathway networks for miR-135b-5p.

## Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SupplementalFigures12.pdf](#)
- [TableS1TumorcharacteristicsinTCGAcohort.docx](#)
- [TableS2DMLs.xlsx](#)
- [TableS3methmiRcorrelations.xlsx](#)
- [TableS4correlationvalidationinTCGA.xlsx](#)
- [TableS5enrichmentanalysis.xlsx](#)